Evaluation of integrating SAP NetWeaver™ based systems in Apache Kafka®

INIT.

ODP Source Connector
Hightlights

AIMS

- Continuous CDC-based data extraction from SAP NetWeaver™ based systems to Apache Kafka®

APPROACH

- Use of the SAP Operational Data Provisioning interface technology
- Native integration through implementation of the Kafka Connect Connector API

RESULT

- Direct use of a comprehensive pool of business content data sources
- Incremental UI-based configuration with input assistance and validation
- Exactly-once data delivery
- Graceful back-off strategy
- Schema evolution compatibility

Motivation
"Einfach machen..."

Due to its high degree of scalability, reliability and failure safety, the **Apache Kafka streaming platform** enjoys increasing popularity for designing data flows in heterogeneous integration architectures. More and more companies wish to integrate this platform seamlessly into their existing IT infrastructure.

There are various approaches for connecting SAP systems like the usage of classical ETL frameworks or cloud-based products such as SAP Data Intelligence. In general, such systems cause high costs in terms of implementation and maintenance. Therefore, it is advisable to examine individually how economical the wide range of functions of such a system is for the company. On the other hand, there are lean solutions that can be integrated natively into the Kafka platform, have no unnecessary platform dependencies, and solely focus on data exchange.

As part of an evaluation project, Mercedes-Benz Bank AG gave us the opportunity to demonstrate our approach of integrating SAP NetWeaver™ based systems in Apache Kafka.

Approach
Kafka Connect

With the Connect component, Kafka offers a native framework for linking various source systems to Apache Kafka. Despite an extensive selection of connectors, there is currently just a HANA connector for interaction with SAP systems available which only allows a direct connection at the database level. Bypassing the application level prevents the utilization of services and functions of the source system and is therefore considered an anti-pattern.
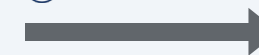
Solution
INIT ODP Source Connector

Our in-depth expertise in technologies related to SAP and Apache Kafka enabled us to develop the native **INIT ODP Source Connector**, which overcomes all previously mentioned challenges. It allows the extraction and replication of various SAP data sources to Apache Kafka, only through configuration adjustments, without additional implementation effort and regardless of the selected database technology. The technical infrastructure is provided by CONFLUENT® and SAP with the CONFLUENT platform and **O**perational **D**ata **P**rovisioning.

SAP SAP NetWeaver → Kafka Connect → CONFLUENT

Contact person: ronald.schertle@init-software.de

Evaluation of integrating SAP NetWeaver™ based systems in Apache Kafka®

Result
Functionality and usability

By using SAP ODP and the associated availability of numerous business content data sources as well as the support of customer-specific extractors, we guarantee a wide range of applications for the connector.

In combination with the CONFLUENT platform, the lean **INIT ODP Source Connector** can be installed, configured, and managed effortlessly. The configuration in the Control Center UI benefits from the ad-hoc value recommendations of the built-in **metadata extractor** and the intensive validation mechanisms of the connector. In addition to continuous delta loads, there are three modes for the initial loading of historical data, which can be adapted to the individual needs using projections and selections. Besides, the metadata extractor ensures automated **schema generation** and data conversion into the internal Kafka Connect format. That allows the **use of existing Kafka Connect Converters** to provide the data in different formats such as Avro or JSON.

In conjunction with the Connector API and the request-based confirmation mechanism in ODP, the connector-specific **offset handling** guarantees **exactly once semantics** for the data delivery. The required information is persisted both in the source system and in Kafka so that connection failure or temporary downtimes of the source system or the Connect platform do not change this guarantee. In addition, a built-in **back-off strategy** prevents utilization peaks in all participating systems in the event of temporary connection problems.

Conclusion
Successful cooperative evaluation

Within this evaluation project with Mercedes-Benz Bank AG, the INIT ODP Source Connector was successfully used to extract financial accounting data from an SAP FI system to support various use cases in the area of bank management and advanced analytics. Performance tests confirmed an extraction rate of **2,000 messages per second**. During an initial load, the whole process of processing in SAP, extracting, converting into Avro, loading into a Kafka topic and message commitment by Kafka took less than 17 minutes of runtime in total. In all executed tests, the connector exceeded every expectation in terms of performance, maintainability, and reliability and convinced everyone involved of its straightforward and universal applicability.
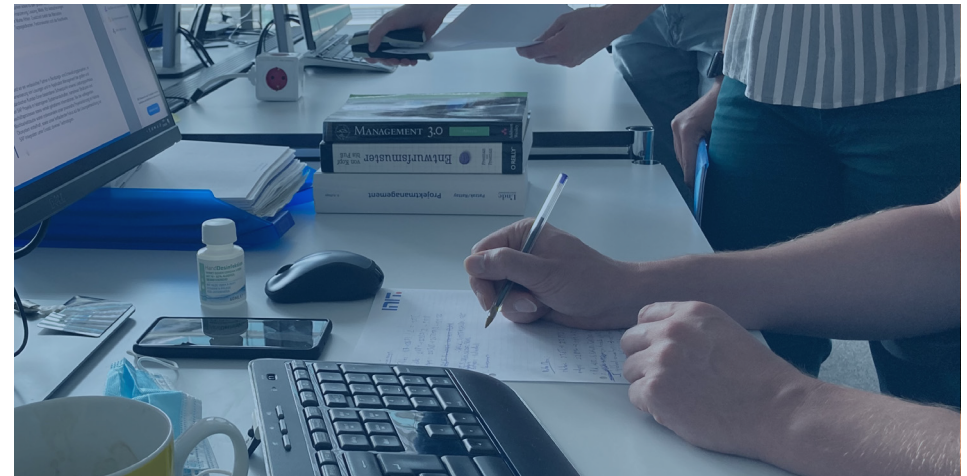
INIT Individuelle Softwareentwicklung & Beratung GmbH

A central focus of the service portfolio of INIT GmbH is SAP projects in heterogeneous system landscapes, complex structures, and business processes as well as distributed information. In the present solution evaluation, our practical project experience in the Hadoop ecosystem was particularly advantageous, as well as our ongoing focus on developing solutions for SAP integration using various technologies.

Scan this code to download the CONFLUENT gold verified INIT ODP Source Connector
confluent.io/hub/init/kafka-connect-odp